Inbreeding Depression in Small Populations of Self-Incompatible Plants

Sylvain Glémin,*,† Thomas Bataillon,* Joëlle Ronfort,* Agnès Mignot† and Isabelle Olivieri†

**INRA-SGAP Montpellier, Domaine de Melgueil, F-34130 Mauguio, France and †Génétique et Environnement CC065, Institut des Sciences de l'Evolution, Université Montpellier II, F-34095 Montpellier, France*

Manuscript received August 8, 2000
Accepted for publication August 10, 2001

ABSTRACT

Self-incompatibility (SI) is a widespread mechanism that prevents inbreeding in flowering plants. In many species, SI is controlled by a single locus (the S locus) where numerous alleles are maintained by negative frequency-dependent selection. Inbreeding depression, the decline in fitness of selfed individuals compared to outcrossed ones, is an essential factor in the evolution of SI systems. Conversely, breeding systems influence levels of inbreeding depression. Little is known about the joint effect of SI and drift on inbreeding depression. Here we studied, using a two-locus model, the effect of SI (frequency-dependent selection) on a locus subject to recurrent deleterious mutations causing inbreeding depression. Simulations were performed to assess the effect of population size and linkage between the two loci on the level of inbreeding depression and genetic load. We show that the sheltering of deleterious alleles linked to the S locus strengthens inbreeding depression in small populations. We discuss the implications of our results for the evolution of SI systems.

SELF-INCOMPATIBILITY (SI) systems provide an effective mechanism for preventing self-fertilization and have evolved repeatedly among flowering plants. Most systems employ physiological mechanisms that prevent pollen germination or pollen tube growth when incompatibility phenotypes are shared by pollen and pistil. Incompatibility processes involve a recognition step, with complex interactions between pollen and pistil that induce acceptance or rejection of the pollen (De Nettancourt 1997). Molecules involved in recognition are usually encoded by genes of one locus (the S locus) with numerous alleles that are maintained by negative frequency-dependent selection (Wright 1939). Gametophytic self-incompatibility systems (GSI) are the most common and they have been noted in 60 families (Kao and McCubbin 1996), whereas sporophytic self-incompatibility systems (SSI) have been observed in only a few families (at least 8; Weller *et al.* 1995). The origin of SI systems in angiosperms is a debated question and different models have been proposed to explain their evolution from self-compatible plants (Charlesworth and Charlesworth 1979; Charlesworth 1988; Uyenoyama 1988a,b,c, 1989, 1991). As in classical models of evolution of selfing and outcrossing initiated by Fisher (1941), these models balance inbreeding depression and the "cost of outcrossing." Interactions between viability loci, causing inbreeding depression, and the S locus are also crucial (see Uyenoyama 1988a,b,c, 1989, 1991). Once SI has evolved, self-compatibility (SC) can secondarily evolve through the breakdown of a SI system if inbreeding depression is small enough (Charlesworth and Charlesworth 1979). Some cases have been documented for GSI systems (Goodwillie 1999), for SSI ones (Reinartz and Les 1994), and for heteromorphic incompatibility (Barrett *et al.* 1989). In large populations, shifts in the level of inbreeding depression are "locked" because of its coevolution with breeding systems: Self-incompatibility and more generally allogamy prevent the purging of deleterious alleles causing inbreeding depression. Therefore, SI systems should not break down in large populations. On the contrary, a recent theoretical study showed that weak inbreeding depression is expected in small populations (Bataillon and Kirkpatrick 2000). This could allow breakdown of SI more easily in accordance with different observations of the dissolution of SI systems after repeated bottlenecks or founder events (*e.g.*, Barrett *et al.* 1989; Reinartz and Les 1994). Moreover, in small populations, because the number of S alleles maintained can be small, the "proportion of plants in a population that are compatible to a given gamete of a randomly chosen individual" (mate availability, Vekemans *et al.* 1998, p. 21) will be low and this could select for SC mechanisms, providing reproductive assurance (Baker 1955, 1967). Selection for reproductive assurance is thus another pressure that might lead to breakdown of SI systems in small populations or in populations that experienced repeated bottlenecks. However, other authors have not found any evidence for the evolution of SC in species that present such conditions (Whishler and Snow 1992; Colas *et al.* 1997) and a comparison of self-incompatibility levels in 11 genera showed no consistent differ-

Corresponding author: Sylvain Glémin, Station de Génétique et d'Amélioration des Plantes INRA Montpellier, Domaine de Melgueil, F- 34130 Mauguio, France. E-mail: glemin@isem.univ-montp2.fr

ences between restricted species and their more widespread congeners (Karron 1987).

Inbreeding depression is a key parameter for the understanding of the evolution of SI systems in small populations. Previous studies dealing with inbreeding depression took into account interactions between viability loci and the S locus (see Uyenoyama 1988a,b,c, 1989, 1991) or finite population size (Bataillon and Kirkpatrick 2000) but never both. Moreover, one can expect the effect of balancing selection at the S locus on other loci, at least those that are tightly linked, to be important in small populations where linkage disequilibrium caused by genetic drift can be rather high (Ohta and Kimura 1969).

The goal of this article is to determine the expected level of inbreeding depression in a population of self-incompatible plants, *i.e.*, to analyze explicitly the effect of the S locus on genes causing inbreeding depression. Here, we present a two-locus model with a GSI locus controlling the mating system and a locus subject to recurrent deleterious mutations that would cause inbreeding depression upon the suppression of SI. Because the interactions between balancing selection and purifying selection in finite populations can be complex, we restricted our study to GSI systems, which are simpler and more common than SSI ones. On one hand, we considered both mildly deleterious and partially recessive mutations and highly recessive lethal mutations, as experimental studies show that deleterious mutations causing inbreeding depression fall into these main categories (*e.g.*, Simmons and Crow 1977). On the other hand, we particularly focused on the effect on expected levels of inbreeding depression of the linkage between the two loci and the population size.

METHODS

Presentation of the model: We studied a population of N diploid hermaphroditic plants with a GSI system. For each individual, we consider two loci: the S locus, which controls crossing between individuals, and another locus, denoted A, undergoing mutation and viability selection and causing inbreeding depression. Under GSI, at least three alleles must segregate and all individuals are heterozygous at the S locus. Crosses between two individuals are allowed only if the allele carried by the pollen grain is different from both the alleles carried by the style; otherwise, crosses are incompatible. No other selection occurs at the S locus. New alleles are generated with rate μ_S, according to the infinite-alleles model (Kimura and Crow 1964). The A locus is diallelic, with a wild-type allele, denoted A, and a partially recessive and deleterious allele, denoted a. Mutations occur at rate μ_1 from A to a and at rate μ_2 (with $\mu_2 \ll \mu_1$) from a to A. Relative fitnesses of genotypes AA, Aa, and aa are, respectively, 1, $1 - hs$, and $1 - s$ ($0 < s \leq 1$ and $0 < h < 1/2$), where s is the selection coefficient and h is the dominance coefficient. Recombination between the two loci occurs at rate r.

Analytical model for a neutral gene linked to the S locus: To understand the effect of the S locus on linked loci, we first investigated the effect of the S locus on a linked neutral locus. To do this, we determined analytically the value of F_{IS}, which expresses the deficit or excess of heterozygotes, due to nonrandom union of gametes, compared to the panmictic case: $F_{IS} = (H_e - H_o)/H_e$, where H_e is the heterozygosity expected under random union of gametes and H_o is the observed heterozygosity. We used this parameter as a quantitative indicator of the effect of the S locus on the A locus. Its value gives insight into the fate of deleterious alleles, as we expect deleterious alleles to be more efficiently selected against in aa genotypes than in Aa genotypes where they are sheltered. For example, in an infinitely large population, the expected equilibrium frequency of a deleterious allele is $\mu_1/s(h + F_{IS} - hF_{IS})$, where F_{IS} holds for a neutral locus (see, for example, Crow and Kimura 1970). At the S locus, noting that homozygotes are absent ($H_o = 1$), and that $H_e = 1 - 1/n_e$ (where n_e is the effective number of S alleles), we have

$$F_{IS}^{S} = -\frac{1}{n_e - 1}.$$

To compute the F_{IS} expected at a neutral locus linked to the S locus, we use a coalescent method developed by Takahata and Satta (1998) for human histocompatibility system (HLA) loci. As in SI systems, numerous alleles are maintained at HLA loci by balancing selection and their coalescent properties are similar to those of GSI loci. Moreover, we can use mean coalescence time to compute F-statistics (Slatkin 1991). Using these properties (see appendix a), we obtain

$$F_{IS} \approx -\frac{2f_S(n-1)}{n^2 + 2f_S(n-1)(4Nr + n - 1)}, \qquad (1)$$

where N is the population size, n is the number of common S alleles (which is approximately equal to n_e, the effective number of S alleles; see Takahata 1990; Vekemans and Slatkin 1994), r is the recombination rate, and f_S is the scaling factor of the genealogy of S alleles (Vekemans and Slatkin 1994). Allelic genealogy is the genealogy of functionally distinct alleles sampled from a finite population and it is similar to a neutral gene genealogy but with a larger time scale, which is equivalent to changing the effective population size from N to Nf_S (Takahata 1990). f_S is always >1 and increases with decreasing N and μ_S. F_{IS} is an increasing function of r. For a large range of parameter values (especially small μ_S and small N), a useful approximation of F_{IS} is

$$F_{IS} = -\frac{1}{n + 1 + 4Nr} \qquad (2)$$

(see appendix a).

Simulation methods: We simulated a population of N diploid individuals with nonoverlapping generations. In each generation, one individual is drawn at random as a mother. After recombination, which occurs with probability r (random sampling in a Bernoulli distribution), one ovule is chosen. A second individual is randomly drawn as the father. After recombination, one pollen grain is chosen. If the S allele carried by the pollen is different from both of the two alleles carried by the mother, a new zygote is formed. If not, the pollen is discarded (incompatible cross) and a new father is drawn. This procedure is iterated until a compatible cross is found (pollen is not limiting). We thus assume no "fecundity selection" (Vekemans *et al.* 1998) to investigate the expected level of inbreeding depression independently of mate availability. Then, the new zygote undergoes selection. A random number is drawn from a uniform distribution on [0, 1]; if the fitness of the zygote (determined by the A locus) is higher or equal to this number, the zygote is kept; if not, it is discarded and the procedure starts again by sampling a mother. Reproduction and selection procedures are iterated until N new zygotes are formed. Mutations at each locus are then applied to the zygotes. At the S locus, the number of mutants is drawn from a Poisson distribution with mean $2N\mu_S$ and then applied at random to genes in the zygotes. At the A locus, for each allele of each individual, a mutant is formed with a probability μ_1 if the allele is A and a probability μ_2 if the allele is a.

The simulation process begins with $2N$ different alleles at the S locus (as in Schierup *et al.* 1997) and alleles A and a are drawn at random from a binomial distribution with mean ½. This initial condition is used in the neutral ($s = 0$) and in the selective case ($s > 0$). To test the sensitivity to the initial conditions, for $s > 0$, simulations were performed with the deleterious alleles initially absent. Results were the same (data not shown). For the range of parameters we used, preliminary computer simulations showed that the mean and variance of the monitored variables became quite stable after ~20,000 generations. Conservatively, we assume that mutation-selection-drift equilibrium was reached after 50,000 generations. At this time we started to record varied information (see later), every 3000 generations for 20 cycles. For each set of parameters, 50 runs were performed to give 1000 values.

To obtain the frequency distributions of mildly deleterious alleles, additional simulations were done to give 200,000 values.

For the A locus alone, the program was checked both with and without selection against expectations for mutation-(selection) drift balance derived from diffusion equations for $N > 100$ (Crow and Kimura 1970) and against expectations derived from a Wright-Fisher matrix model for $N < 100$. For the S locus alone, the program was checked against expectations derived from diffusion equations (for $N > 100$; Yokoyama and Nei 1979) and against published simulation results (Schierup *et al.* 1997).

The stochastic simulation program was developed in Turbo Pascal language (Delphi 4).

Monitored variables: Inbreeding depression, δ, is defined as the decline in fitness of individuals produced by selfing (W_S) relative to the fitness of outcrossed individuals (W_O) (Charlesworth and Charlesworth 1987). Mutation load, L, is defined as the decline in mean fitness ($\overline{W}$) relative to the fitness of the optimal genotype of the population (Crow and Kimura 1970),

$$\delta = 1 - \frac{W_S}{W_O} \tag{3}$$

and

$$L = 1 - \overline{W}. \tag{4}$$

For the A locus, we determined the frequency of allele a (p_a; mean and variance among runs), F_{IS} (mean and variance), the probability of fixation of both alleles, the level of inbreeding depression (δ; mean and variance), and the mutation load (L; mean and variance) when $s > 0$. W_O and $\overline{W}$ were considered equal and were computed by summing up the frequency of the genotypes, weighted by their fitness, $W_O = \overline{W} = f(AA) \times 1 + f(Aa) \times (1 - hs) + f(aa) \times (1 - s)$; W_S was computed in a similar way, using the frequency of the three genotypes expected after a self-fertilization event, $W_S = f(AA) \times 1 + f(Aa) \times [¼ \times 1 + ½ \times (1 - hs) + ¼ \times (1 - s)] + f(aa) \times (1 - s)$. δ and L were then computed using (3) and (4), respectively. L corresponds to the load in an SI population and δ to the level of inbreeding that would be expressed if SI were suppressed. To compute F_{IS}, H_o was recorded in the simulation and H_e (the heterozygosity expected under random union of gametes in a population consisting of N individuals) was calculated as $H_e = 2p_a (1 - p_a)\, 2N/(2N - 1)$ (*e.g.*, Gale 1990).

For the S locus, we determined the mean and variance of the actual and effective numbers of S alleles and the frequency of S alleles.

Parameter values investigated: Because the aim of this study was to investigate the effects of linkage and population size, other parameters were less fully explored. To modify the number of S alleles without affecting the population size, three mutation rates were used for the S locus ($\mu_S = 10^{-3}$, 10^{-4}, and 10^{-5}). These rates are of the same order as the rates used by Schierup *et al.* (1997) or Vekemans *et al.* (1998). We used rather large mutation rates at the A locus ($\mu_1 = \mu_2 = 10^{-3}$ for $s = 0$; $\mu_1 = 10^{-3}$ and $\mu_2 = 10^{-4}$ for $s > 0$) to maintain enough polymorphism even in very small populations ($N < 50$). Values assigned to s (0.1) and h (0.2) are of the same order as those used in various other studies (*e.g.*, Charlesworth *et al.* 1990, 1992) and compatible with available experimental data. For example, $0.075 < s < 0.105$ in *Drosophila* sp. (reviewed in Simmons and Crow 1977) and $s = 0.2$ in wheat (T. Bataillon, un-

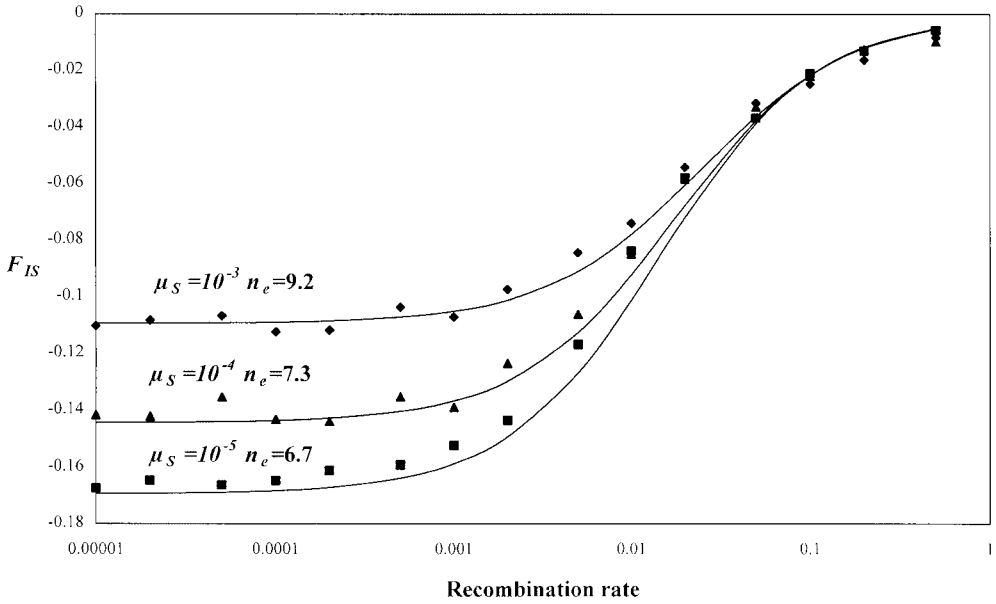

FIGURE 1.—Expected F_{IS} as a function of linkage to the S locus. (N = 100, $\mu_1 = \mu_2 = 10^{-3}$, and $\mu_S = 10^{-5}$, 10^{-4}, and 10^{-3}, giving, respectively, n_e = 6.3, 7.3, and 9.2.) Continuous curves are given by Equation 1 and plotted symbols represent mean results of 1000 simulations.

published data); $0.1 < h < 0.4$ in *Drosophila* sp. (reviewed in SIMMONS and CROW 1977) and in two species of Amsinckia (Boraginaceae; JOHNSTON and SCHOEN 1995) and $0.09 < h < 0.21$ in *Mimulus guttatus* (WILLIS 1999). We also tested the case of almost completely recessive lethal alleles, assuming the parameter values $h = 0.02$ and $s = 1$.

RESULTS

Effect of linkage to the S locus: To investigate the effect of linkage between the two loci we considered a single population size. The results below are for $N = 100$ (simulations run for different population sizes showed essentially the same pattern, results not shown).

F_{IS} for a neutral locus as a function of its linkage to the S locus: Figure 1 shows that for loci that are linked to the S locus, F_{IS} is negative and decreases as the recombination rate decreases. Inspection of Equation 1 shows that this behavior is expected whatever the value of the parameters (see APPENDIX A). Varying the mutation rate (μ_S) allows changing the number of S alleles maintained in the population without modifying the population size. The mutation rate also affects the scaling factor, f_S, but for the parameters we used (N and μ), F_{IS} is quasi-independent of f_S (see approximation in Equation 2). At linked loci, the excess of heterozygotes ($F_{IS} < 0$) is greater for low mutation rates, *i.e.*, for few S alleles maintained in the population (Figure 1 and Equation 2). As r tends toward 0, F_{IS} tends toward $-1/(n-1) \approx F_{IS}^S = -1/(n_e - 1)$, the expected F_{IS} at the S locus (see Equation 2). For $r > 0.1$, the S locus has a negligible effect on F_{IS}.

Effect of linkage to the S locus on selection against deleterious alleles: When the S and A loci are linked, the distribution of the deleterious allele frequency is greatly modified compared to the case of finite panmictic populations (without SI) or to the case of an unlinked locus (Table 1). The mean frequency of the deleterious allele strongly increases as the recombination rate decreases (from 0.04 for an unlinked locus to 0.32 for $r = 10^{-4}$, with $\mu_S = 10^{-5}$; Table 1) and the fixation probability decreases with increasing linkage (from 0.34 to 0 for the same values; Table 1). Consequently, for tightly linked loci, the frequency distribution is much less skewed than for unlinked loci in SI populations or for loci in panmictic (SC) populations (see Figure 2 for comparison with the panmictic case). As for the neutral case, the smaller the number of S alleles (*i.e.*, the smaller the mutation rate, μ_S), the stronger the effect on the linked locus: The mean frequency of deleterious alleles is higher (see Table 1) and the frequency distribution is less skewed (see Figure 2).

Consequences for inbreeding depression and the mutation load: Both mean inbreeding depression and mean mutation load (Table 1) strongly increase as the recombination rate decreases. For fixed recombination rates, both mean inbreeding depression and mean mutation load are higher for smaller numbers of S alleles maintained in the population (compare values across mutation rates in Table 1). We observed the same pattern of variation for a lethal gene (data not shown).

Effect of the A locus on the S locus: Balancing selection at the S locus modifies the distribution of allele frequencies at linked loci. Reciprocally, one might expect that selection against deleterious alleles modifies the allelic dynamics at the S locus. However, even in the case of a lethal gene, purifying selection among alleles linked to the S locus has negligible effects on S alleles' dynamics (number and distribution of S alleles; data not shown).

Effect of population size: We studied the effect of population size for different values of recombination

TABLE 1

Effect of linkage to the S locus on partially recessive and mildly deleterious alleles

	Panmixia					
	0.041 (0.058)			0.317		
	$\mu_S = 10^{-3}$	$\mu_S = 10^{-4}$	$\mu_S = 10^{-5}$	$\mu_S = 10^{-3}$	$\mu_S = 10^{-4}$	$\mu_S = 10^{-5}$
	Allele frequency			Probability of fixation		
$r = 0.5$	0.040 (0.058)	0.038 (0.055)	0.042 (0.061)	0.316	0.328	0.32
$r = 0.1$	0.042 (0.059)	0.043 (0.061)	0.042 (0.062)	0.298	0.307	0.306
$r = 0.01$	0.073 (0.082)	0.073 (0.083)	0.072 (0.088)	0.213	0.199	0.228
$r = 0.001$	0.114 (0.101)	0.170 (0.124)	0.191 (0.130)	0.123	0.076	0.067
$r = 0.0001$	0.116 (0.102)	0.247 (0.130)	0.323 (0.137)	0.123	0.024	0.007
$r = 0.00001$	0.123 (0.104)	0.259 (0.137)	0.365 (0.144)	0.118	0.017	0.004
$r = 0$	0.133 (0.105)	0.255 (0.135)	0.369 (0.139)	0.102	0.023	0.002
	Panmixia					
	1.07 (1.38)			1.93 (3.12)		
	$\mu_S = 10^{-3}$	$\mu_S = 10^{-4}$	$\mu_S = 10^{-5}$	$\mu_S = 10^{-3}$	$\mu_S = 10^{-4}$	$\mu_S = 10^{-5}$
	Inbreeding depression (×1000)			Mutation load (×1000)		
$r = 0.5$	1.09 (1.45)	1.03 (1.38)	1.13 (1.49)	1.86 (3.07)	1.78 (2.84)	1.97 (3.21)
$r = 0.1$	1.14 (1.48)	1.17 (1.52)	1.14 (1.53)	1.96 (3.11)	2.02 (3.18)	1.95 (3.23)
$r = 0.01$	1.97 (2.05)	1.98 (2.10)	1.95 (2.20)	3.38 (4.29)	3.34 (4.25)	3.36 (4.60)
$r = 0.001$	3.05 (2.41)	4.32 (2.69)	4.85 (2.78)	5.39 (5.58)	8.47 (7.53)	9.49 (8.05)
$r = 0.0001$	3.09 (2.45)	5.97 (2.45)	7.26 (2.06)	5.47 (5.67)	13.0 (8.78)	18.1 (10.7)
$r = 0.00001$	3.27 (2.46)	6.10 (2.44)	7.69 (1.86)	5.86 (5.82)	13.9 (9.62)	21.5 (12.1)
$r = 0$	3.50 (2.45)	6.03 (2.36)	7.81 (1.77)	6.35 (5.90)	13.6 (9.63)	21.7 (11.5)

Frequency of the partially recessive and mildly deleterious alleles ($h = 0.2$, $s = 0.1$, $\mu_1 = 10^{-3}$, and $\mu_2 = 10^{-4}$), probability of fixation, inbreeding depression, and the mutation load are given for a panmictic population without SI (panmixia) and for the A locus linked to the S locus with different recombination rates (r). For the second case (linkage), three mutation rates at the S locus are presented. $N = 100$ and the effective numbers of S alleles maintained in the population are $n_e = 6.7$ for $\mu_S = 10^{-5}$, $n_e = 7.3$ for $\mu_S = 10^{-4}$, and $n_e = 9.2$ for $\mu_S = 10^{-3}$. Means and standard deviations are calculated over 1000 replicates.

rates ($r = 0.5$, 10^{-2}, 10^{-3}, and 10^{-4}). For neutral loci linked to the S locus, the smaller the population size, the more negative F_{IS} (Figure 3). Heterozygosity is therefore larger in small populations and for loci that are more tightly linked to the S locus. This result can be directly deduced from Equation 2: First, F_{IS} is inversely related to N; second, it is also inversely related to n and, for a given value of μ_S, fewer S alleles are maintained in small populations than in large ones.

Frequency of deleterious alleles as a function of population size: As in the previous section, our simulation results without SI or with free recombination between the two loci agree with results for finite panmictic populations. The mean frequency of partially recessive and mildly deleterious alleles weakly decreases as population size decreases except for very small populations where it increases (Table 2). For lethal alleles (Table 3), mean frequency uniformly decreases as population size decreases (*e.g.*, Crow and Kimura 1970). On the contrary, when loci are (closely) linked to the S locus, the mean frequency of both deleterious and lethal alleles strongly increases with decreasing population size (Tables 2 and 3). Figures 4 and 5 show the excess of deleterious or lethal alleles caused by linkage to the S locus compared to a panmictic population (without SI), expressed as the ratio (R) of the mean frequencies. R can be very high (12 for a mildly deleterious allele, 18 for a lethal allele). Moreover, this excess is maximal for intermediate or small population sizes (in our simulations, $N = 50$ for mildly deleterious alleles; $N = 25$ and probably less for lethal alleles: No simulation was run for $N < 25$ because the minimum of three S alleles could not be maintained). In addition, even in small populations, fixation probabilities are small for tight linkage (data not shown).

Inbreeding depression and mutation load as a function of population size: As previously, our simulation results without SI and for $r = 0.5$ agree with results for panmictic populations. The mean mutation load due to partially recessive and mildly deleterious alleles increases as population size decreases (Table 2); for recessive lethal alleles (Table 3), the load is insensitive to population size, except in very small populations where it decreases (*e.g.*, Crow and Kimura 1970). For mildly dele-

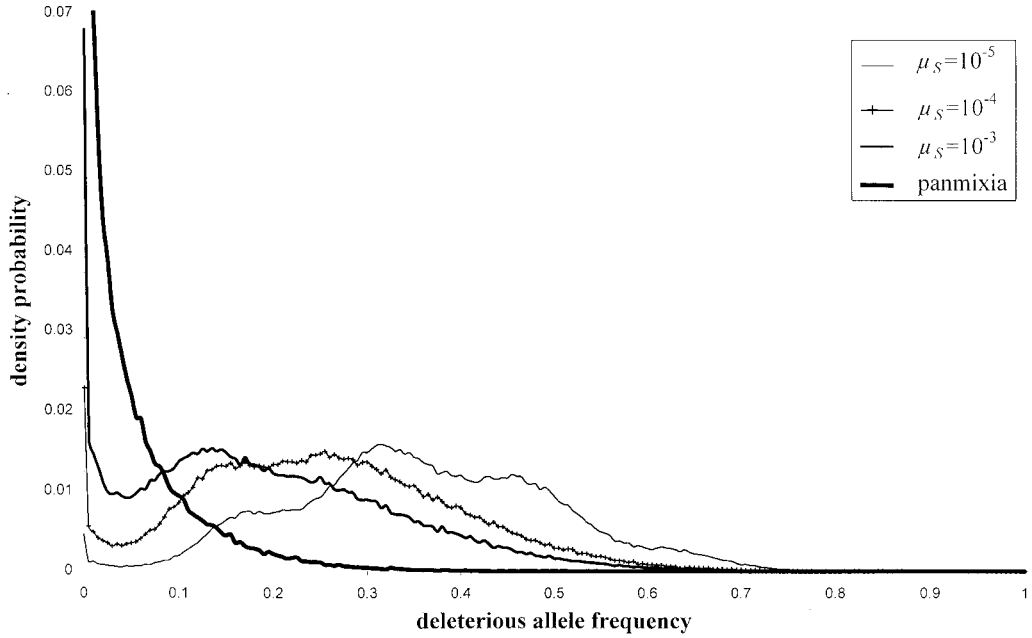

FIGURE 2.—Probability distribution of deleterious allele frequency based on 200,000 simulations. ▬, frequency distribution of a recessive mildly deleterious allele in a panmictic population ($N = 100$, $h = 0.2$, $s = 0.1$, $\mu_1 = 10^{-3}$, $\mu_2 = 10^{-4}$). Frequency distribution of a recessive mildly deleterious allele at a locus closely linked ($r = 10^{-4}$) to the S locus ($N = 100$, $h = 0.2$, $s = 0.1$, $\mu_1 = 10^{-3}$, $\mu_2 = 10^{-4}$). Mutation rates assumed for the S locus are as follows: ——, $\mu_S = 10^{-3}$ (giving $n_e = 9.2$); +—, $\mu_S = 10^{-4}$ (giving $n_e = 7.3$); —, $\mu_S = 10^{-5}$ (giving $n_e = 6.7$).

terious alleles linked to the S locus, the mutation load strongly increases in small populations, much more than in the panmictic case (Table 2). For lethal alleles linked to the S locus, the load also strongly increases in small populations, contrary to what is expected in a panmictic population (Table 3).

For inbreeding depression, patterns are very different between the cases with or without linkage. In a panmictic population or with no linkage between the two loci, inbreeding depression decreases when population size decreases (Tables 2 and 3). On the contrary, for deleterious alleles linked to the S locus, inbreeding depression is higher in small populations than in large ones (Tables 2 and 3).

DISCUSSION

Effect of the S locus on linked loci: Selection at a locus can modify selection dynamics at linked loci. Numerous studies have dealt with the effect of selection on neutral linked polymorphism (*e.g.*, CHARLESWORTH *et al.* 1997; TAKAHATA and SATTA 1998). However, the effect of a selected locus on another selected one has not been much investigated. In our case, in addition to local ef-

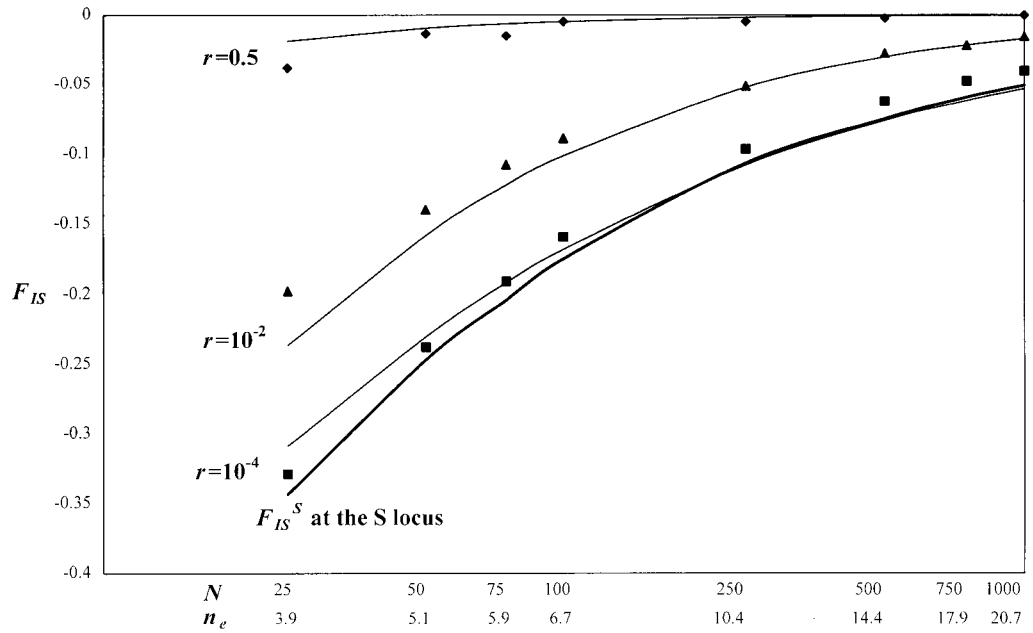

FIGURE 3.—Expected F_{IS} for a neutral locus as a function of population size, N, and of the effective number of S alleles, n_e, maintained in the population, for different recombination rates between the S and the A locus. Mutation rates for all cases are $\mu_1 = \mu_2 = 10^{-3}$ and $\mu_S = 10^{-5}$. Continuous curves are given by Equation 1 and plotted symbols represent mean results of 1000 simulations. F_{IS}^{S} at the S locus itself is also presented (thick line).

TABLE 2

Effect of genetic drift and linkage on partially recessive and mildly deleterious alleles

N	n_e	Panmixia	$r = 0.5$	$r = 10^{-2}$	$r = 10^{-3}$	$r = 10^{-4}$
				Allele frequency		
25	3.9	0.101 (0.28)	0.104 (0.261)	0.183 (0.293)	0.453 (0.304)	0.629 (0.241)
50	5.1	0.038 (0.077)	0.042 (0.081)	0.098 (0.135)	0.290 (0.191)	0.462 (0.191)
75	5.9	0.036 (0.063)	0.041 (0.067)	0.078 (0.098)	0.219 (0.148)	0.377 (0.161)
100	6.7	0.040 (0.059)	0.040 (0.053)	0.071 (0.084)	0.196 (0.131)	0.331 (0.141)
250	10.4	0.040 (0.039)	0.042 (0.039)	0.057 (0.048)	0.113 (0.074)	0.194 (0.085)
500	14.4	0.042 (0.028)	0.041 (0.028)	0.052 (0.036)	0.080 (0.047)	0.122 (0.057)
750	17.9	0.043 (0.22)	0.043 (0.024)	0.051 (0.028)	0.067 (0.036)	0.090 (0.044)
1000	20.7	0.043 (0.020)	0.043 (0.022)	0.049 (0.024)	0.061 (0.028)	0.074 (0.033)
				Inbreeding depression (×1000)		
25	3.9	0.778 (1.85)	0.809 (1.95)	2.42 (3.26)	6.52 (4.25)	7.49 (3.85)
50	5.1	0.941 (1.73)	1.07 (1.83)	2.47 (3.07)	6.50 (3.26)	8.37 (2.33)
75	5.9	0.932 (1.50)	1.07 (1.60)	2.08 (2.44)	5.42 (3.06)	7.85 (2.10)
100	6.7	1.07 (1.44)	1.09 (1.35)	1.93 (2.13)	4.96 (2.82)	7.32 (2.14)
250	10.4	1.11 (1.03)	1.17 (1.00)	1.62 (1.28)	3.10 (1.84)	5.01 (1.82)
500	14.4	1.19 (0.77)	1.17 (0.75)	1.47 (0.95)	2.26 (1.24)	3.33 (1.39)
750	17.9	1.21 (0.64)	1.22 (0.66)	1.46 (0.76)	1.88 (0.95)	2.50 (1.11)
1000	20.7	1.23 (0.56)	1.23 (0.58)	1.41 (0.65)	1.72 (0.77)	2.08 (0.88)
				Mutation load (×1000)		
25	3.9	8.60 (25.1)	8.76 (23.4)	13.5 (27.3)	32.6 (31.7)	48.5 (28.6)
50	5.1	1.95 (4.53)	2.12 (4.68)	4.90 (7.96)	16.2 (15.0)	29.9 (18.6)
75	5.9	1.70 (3.50)	1.97 (3.63)	3.63 (5.29)	11.2 (9.5)	22.4 (13.8)
100	6.7	1.91 (3.17)	1.85 (2.68)	3.23 (4.24)	9.79 (8.05)	18.7 (10.9)
250	10.4	1.76 (1.91)	1.85 (1.84)	2.52 (2.30)	5.18 (3.86)	9.52 (5.11)
500	14.4	1.81 (1.33)	1.77 (1.28)	2.24 (1.67)	3.54 (2.29)	5.62 (3.03)
750	17.9	1.83 (1.07)	1.85 (1.14)	2.20 (1.31)	2.90 (1.68)	4.02 (2.12)
1000	20.7	1.85 (0.94)	1.85 (0.97)	2.10 (1.08)	2.62 (1.33)	3.22 (1.57)

Frequency of the partially recessive and mildly deleterious allele ($h = 0.2$, $s = 0.1$, $\mu_1 = 10^{-3}$, and $\mu_2 = 10^{-4}$), inbreeding depression, and the mutation load are given for various population sizes. Mutation rate at the S locus for all cases is $\mu_S = 10^{-5}$. Results are given for a panmictic population without SI (panmixia) and for different recombination rates (r) between the S locus and the A locus. Means and standard deviations are calculated over 1000 replicates.

fects on the genome (around the S locus), interactions between a locus controlling the mating system and a locus involved in inbreeding depression can affect the evolution of the mating system itself.

Linkage to the S locus limits selection against deleterious alleles: It is classically recognized that balancing selection allows the maintenance of a large number of S alleles and maintains an excess of heterozygotes at the S locus (Yokoyama and Nei 1979). In agreement with earlier studies (Ohta and Kimura 1970; Charlesworth *et al.* 1997), our analytical and simulation results show that genes that are linked to a locus under balancing selection (GSI here) will also present an excess of heterozygotes ($F_{IS} < 0$: see Equation 1 and Figure 1). Because counterselection of recessive deleterious alleles is primarily directed against homozygotes (*aa*), F_{IS} is a quantitative indicator of the sheltering of deleterious recessive alleles, caused by the linkage to the S locus. Sheltering of deleterious recessive alleles can also be explained by the dynamics of the S alleles. In small populations, drift leads to fluctuations in S-allele frequency and so to fluctuating intensity of frequency-dependent selection. Because balancing selection at the S locus is much stronger than purifying selection, a rare (respectively a frequent) S allele will be selected for (respectively against) whatever the allele (deleterious or not) it carries. Thus, because of the linkage to the S locus, the A locus is also subject to a kind of "associative overdominance" (Ohta and Kimura 1970), which limits the efficiency of selection against deleterious alleles. For fixed population sizes, the sheltering of deleterious alleles is also stronger when the rate of origination of novel S alleles (μ_S) is small. Two processes explain this result. First, for low mutation rates, the number of S alleles maintained is small so balancing selection is strong and the excess heterozygosity is high (see Equation 2). Second, S alleles are maintained in the population for long evolutionary periods (f_S is high; see Vekemans and Slatkin 1994) so linked deleterious alleles can persist in the population as long as S alleles persist or until a

TABLE 3

Effect of genetic drift and linkage on lethal alleles

N	n_e	Panmixia	$r = 0.5$	$r = 10^{-2}$	$r = 10^{-3}$	$r = 10^{-4}$
				Allele frequency		
25	3.9	0.010 (0.028)	0.011 (0.028)	0.048 (0.089)	0.167 (0.124)	0.231 (0.096)
50	5.1	0.014 (0.026)	0.015 (0.028)	0.037 (0.058)	0.125 (0.096)	0.175 (0.078)
75	5.9	0.013 (0.022)	0.015 (0.024)	0.034 (0.048)	0.101 (0.080)	0.140 (0.069)
100	6.7	0.015 (0.020)	0.016 (0.022)	0.035 (0.044)	0.082 (0.066)	0.117 (0.065)
250	10.4	0.018 (0.014)	0.019 (0.017)	0.027 (0.022)	0.045 (0.036)	0.061 (0.040)
500	14.4	0.020 (0.010)	0.021 (0.010)	0.025 (0.014)	0.031 (0.020)	0.036 (0.022)
750	17.9	0.021 (0.010)	0.021 (0.010)	0.024 (0.010)	0.026 (0.014)	0.029 (0.014)
1000	20.7	0.021 (0.010)	0.021 (0.010)	0.023 (0.010)	0.025 (0.010)	0.026 (0.010)
				Inbreeding depression (×1000)		
25	3.9	4.96 (13.6)	5.04 (14.1)	23.4 (43.4)	81.0 (60.6)	112.3 (46.8)
50	5.1	6.51 (12.4)	7.11 (13.5)	17.8 (28.3)	60.4 (46.8)	84.6 (37.8)
75	5.9	6.11 (10.2)	7.00 (11.6)	16.4 (23.4)	78.8 (3807)	67.6 (33.4)
100	6.7	7.13 (9.93)	7.70 (10.8)	16.8 (20.9)	39.4 (32.3)	56.3 (31.2)
250	10.4	8.65 (7.48)	8.92 (7.73)	12.7 (11.0)	21.8 (17.1)	29.6 (19.3)
500	14.4	9.72 (5.69)	9.92 (5.81)	12.0 (7.11)	15.0 (9.67)	17.4 (11.0)
750	17.9	10.0 (5.05)	10.2 (4.98)	11.7 (5.54)	12.6 (6.30)	13.8 (7.23)
1000	20.7	10.2 (4.49)	10.2 (4.29)	11.26 (4.64)	12.23 (5.55)	12.7 (5.60)
				Mutation load (×1000)		
25	3.9	0.412 (1.13)	0.418 (1.17)	1.93 (3.58)	6.68 (4.99)	9.26 (3.84)
50	5.1	0.541 (1.03)	0.591 (1.12)	1.47 (2.35)	4.99 (3.86)	6.99 (3.11)
75	5.9	0.508 (0.849)	0.582 (0.964)	1.36 (1.94)	4.04 (3.20)	5.60 (2.76)
100	6.7	0.593 (0.825)	0.640 (0.900)	1.39 (1.74)	3.27 (2.67)	4.66 (2.59)
250	10.4	0.720 (0.624)	0.742 (0.640)	1.06 (0.911)	1.82 (1.42)	2.46 (1.60)
500	14.4	0.809 (0.469)	0.826 (0.480)	0.998 (0.592)	1.25 (0.806)	1.45 (0.911)
750	17.9	0.833 (0.424)	0.851 (0.412)	0.974 (0.458)	1.05 (0.520)	1.15 (0.600)
1000	20.7	0.847 (0.374)	0.849 (0.361)	0.938 (0.387)	1.02 (0.458)	1.06 (0.469)

Frequency of the lethal allele ($h = 0.02$, $s = 1$, $\mu_S = 10^{-3}$, and $\mu_S = 10^{-4}$), inbreeding depression, and the mutation load are given for various population sizes. Mutation rate at the S locus for all cases is $\mu_S = 10^{-5}$. Results are given for a panmictic population without SI (panmixia) and for different recombination rates (r) between the S locus and the A locus. Means and standard deviations are calculated over 1000 replicates.

recombination event breaks the linkage. As a result, mutation load and inbreeding depression are increased, compared to what is expected for loci unlinked to the S locus.

Thus, theoretically, deleterious mutations can accumulate near the S locus. In addition, Uyenoyama (1997) suggested that these mutations should be S allele specific. S-specific deleterious alleles were found in *Solanum carolinense* (M. K. Uyenoyama and J. L. Stone, personal communication) and, in *Papaver rhoeas*, there is also evidence for a selected locus involved in dormancy linked to the S locus (Lane and Lawrence 1995). Such specific associations between S alleles and deleterious alleles may reduce further the efficiency of selection against deleterious mutants. Indeed, in a two-allele model, deleterious mutations can be associated with different S alleles (unless $N\mu_1$ is very small) and then counterselected (in homozygous state) if they are brought together through mating. Conversely, under specific association, homozygotes can be produced (and thus eliminated) only following a recombination event between the S locus and the selected locus or upon the generation of new S alleles from existing ones (Uyenoyama 1997).

Effect of population size: In a finite panmictic population the mean frequency of a mildly deleterious allele (x^*) is weakly sensitive to the population size for large populations ($Ns \gg 1$) whereas, in small populations ($Ns < 1$), x^* can be high because deleterious alleles can fix because of genetic drift (Kimura *et al.* 1963; see mean allele frequency in panmixia in Table 2). For a lethal recessive allele, mean frequency monotonically decreases as population size decreases (see Crow and Kimura 1970 and Table 3). For genes tightly linked to the S locus, our study shows that the mean frequencies of both mildly deleterious and lethal alleles behave very differently (see Tables 2 and 3 for $r < 0.5$). The excess of deleterious alleles maintained in the population compared to the panmictic case (ratio R in Figures 4 and 5) allows us to analyze the relative magnitude of the different evolutionary pressure (purifying selection, "associative overdominance," and genetic drift) as a function of population size. For mildly deleterious alleles, R

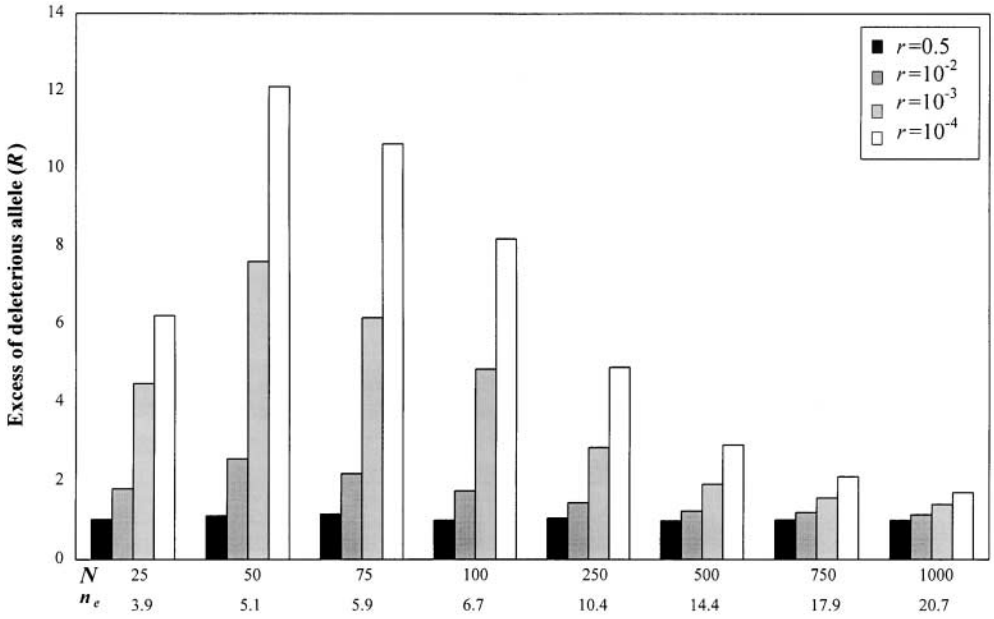

FIGURE 4.—Ratio, R, of the mean frequency of a recessive, mildly deleterious allele linked to the S locus ($r = 10^{-4}$) over its mean frequency in a panmictic population (without SI). R is given as a function of population size, N, and of the effective number of S alleles (n_e) maintained in the population ($h = 0.2$, $s = 0.1$; $\mu_1 = 10^{-3}$, $\mu_2 = 10^{-4}$, $\mu_S = 10^{-5}$).

presents a maximum for intermediate population sizes (Figure 4). In large populations, balancing selection at the S locus is rather weak (n is high) and so is the "associative overdominance" at linked loci, whereas purifying selection against deleterious alleles is strong ($Ns \gg 1$). As in a large panmictic population, purifying selection is the main force determining the frequency of deleterious alleles: R tends toward 1. In very small populations ($N < 50$ in Figure 4), drift is very important and offsets the effect of the S locus on linked genes. So we expect that R also tends toward 1 when N tends toward 0 (but, as explained earlier, simulations could not be done for $N < 25$). For intermediate population sizes, the effect of the S locus becomes very important and linkage to the S locus can induce a great excess of deleterious alleles ($R \approx 12$, Figure 4). For lethal alleles, R is higher (≈ 18 for $N = 25$, Figure 5) and, even in very small populations ($N = 25$), drift does not offset the effect of the S locus.

Another important result of our study is that linkage to the S locus induces higher inbreeding depression in small populations than in large ones (see Tables 2 and 3). With no linkage between the two loci or in a panmictic population without SI, the reverse process is expected. Inbreeding depression decreases as population size decreases (see Tables 2 and 3 and BATAILLON and

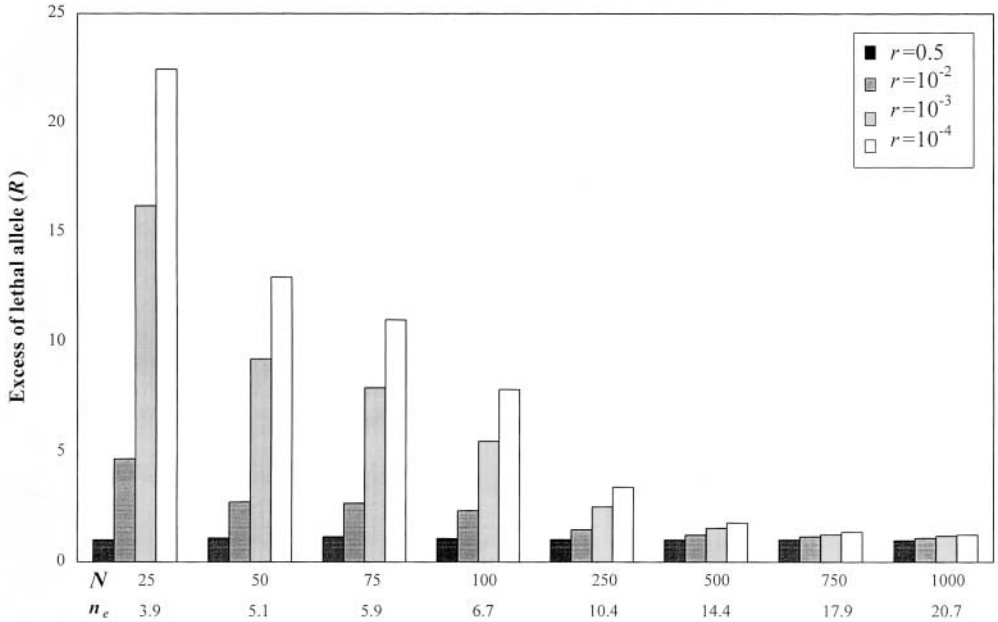

FIGURE 5.—Ratio, R, of the mean frequency of a lethal allele linked to the S locus over its mean frequency in a panmictic population (without SI). R is given as a function of population size, N, and of the effective number of S alleles (n_e) maintained in the population ($h = 0.02$, $s = 1$; $\mu_1 = 10^{-3}$, $\mu_2 = 10^{-4}$, $\mu_S = 10^{-5}$).

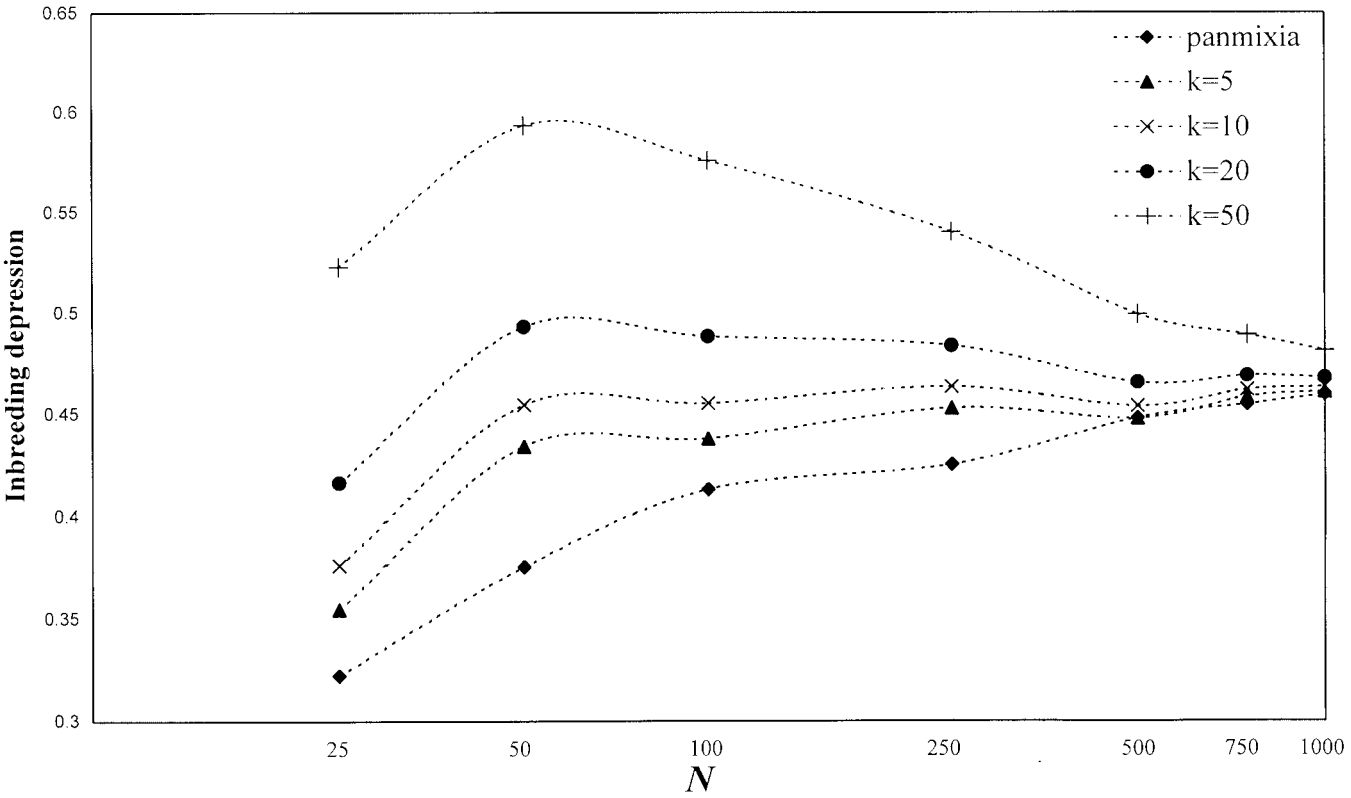

FIGURE 6.—Total inbreeding depression, caused by $p = 500$ loci, as a function of population size. Loci act multiplicatively and k loci are assumed to be linked to the S locus ($r = 10^{-4}$, $\mu_S = 10^{-5}$). The mutation rates toward the deleterious allele are equal to $\mu_1 = 10^{-3}$ ($\mu_2 = 10^{-4}$), giving a genomic mutation rate of $U = p\mu_1 = 0.5$. All the other parameters are assumed to be equal for all loci ($h = 0.2$, $s = 1$). Calculation of total inbreeding depression is given in APPENDIX B.

KIRKPATRICK 2000). This is due to the fact that the probability of fixation of one of the two alleles (A or a) is large for small N and inbreeding depression is absent in a monomorphic population. For genes contributing to inbreeding depression linked to the S locus, we have shown that the mean frequency of the deleterious allele is strongly increased (Tables 2 and 3, Figures 4 and 5). In this case, however, the probability of fixation is now very low (see, for example, Table 1 for $N = 100$) and the A locus is still polymorphic so that inbreeding depression is strongly increased.

Maintenance and evolution of SI systems in small populations: It has been claimed that selection for reproductive assurance could lead to the evolution of SC from SI populations (BAKER 1955, 1967; GOODWILLIE 1999). Limitation of compatible pollen can occur after founder events (BAKER 1955, 1967), in populations experiencing repeated bottleneck (REINARTZ and LES 1994), and more generally in small populations where few S alleles can be maintained (BYERS and MEAGHER 1992; VEKEMANS *et al.* 1998). This can lead to a reduction of the seed set as was observed in self-incompatible and partially clonal populations (where local effective population size and mate availability are small) such as *Podophyllum peltatum* (Berberidaceae; WHISHLER and SNOW 1992) and *Filipendula rubra* (Rosaceae; ASPINWALL and CHRISTIAN 1992). Moreover, because small levels of inbreeding depression are expected in small populations, this could not offset the advantage of SC (BATAILLON and KIRKPATRICK 2000). However, no evidence of evolution of SC was found in *P. peltatum* (WHISHLER and SNOW 1992) and a comparison of self-incompatibility levels in 11 genera showed no consistent differences between restricted species and their more widespread congeners (KARRON 1987). Contrary to the results of BATAILLON and KIRKPATRICK (2000), our model suggests that populations that express GSI can maintain rather strong inbreeding depression, even in small populations. Such strong inbreeding depression could explain the maintenance of SI systems where breakdown would be expected. However, to maintain high levels of inbreeding depression, sufficient numbers of loci must be linked to the S locus. Figure 6 shows the level of inbreeding depression expected as a function of population size for different numbers of loci (k) tightly linked to the S locus. Other unlinked loci ($p - k$) are also involved and a genomic mutation rate of $U = 0.5$ is assumed (see details in APPENDIX B). If k is not too small, rather high levels of inbreeding depression can be maintained for a wide range of population sizes.

However, we do not assume an effect of epistasis between viability loci, which can modify predictions quantitatively.

How many genes can be under the influence of the S locus? The influence of balancing selection at the S locus is limited to a closely linked genomic region, typically for $Nr < 1$ (see Equation 2 and previous discussion). This is consistent with predictions of Strobeck (1980, 1983), who suggested that the effect of a locus under balancing selection on heterozygosity (Strobeck 1980) and linkage disequilibrium (Strobeck 1983) becomes important for $Nr < 1$. The number of genes under the influence of the S locus will depend on the genes density and on the local recombination rates in this genomic region. Some results suggest that this number of genes should be high. The S locus region seems to be a region with low recombination rates in several species. For example, in *Lycopersicum esculentum* (Pillen *et al.* 1996) and in *Petunia hybrida* (Entani *et al.* 1999) the S locus is located in a centromeric region where recombination is usually low. In the Solanaceae, McCubbin and Kao (1999) suggest that the S locus is embedded in a gene-rich region, and Dowd *et al.* (2000) have located 13 pollen-expressed genes in a 1-cM region around the S locus in *P. inflata*. In addition, a recent theoretical study reported that, for a neutral locus, the chromosome region under the influence of a balancing selection locus is wider in subdivided populations than in nonsubdivided ones (Schierup *et al.* 2000). This might reinforce the conclusions of our models where we only considered a single population. However, interaction between balancing selection, purifying selection, and population subdivision is still a very complex question and direct extension of our work to subdivided populations could be misleading.

Extensions to other SI systems: SSI systems are more complex than GSI ones. In codominant SSI systems (SSIcod in Schierup *et al.* 1997), results would be qualitatively similar: A high level of inbreeding depression is expected for loci linked to the S locus because balancing selection is also strong and symmetrical and can maintain high heterozygosity near the S locus. With dominance relationships in both the pollen and the stigma (SSIdom in Schierup *et al.* 1997) or only in the pollen (SSIdomcod in Schierup *et al.* 1997), S-allele dynamics are more complex. Balancing selection is weaker and not symmetrical. While our model cannot offer clear predictions for SSI systems (especially for SSI with dominance relationships), we still expect that an SSI locus, as well as a GSI one, can shelter linked deleterious alleles. In addition, as in the Solanaceae, a rather high number of loci can be under the influence of the S locus in the Brassicaceae. In *Brassica* sp., divergent and rearranged sequences may suppress recombination within the S-locus complex (Boyes *et al.* 1997; but see Awadalla and Charlesworth 1999 and Casselman *et al.* 2000) and Susuki *et al.* (1999) estimated a high gene density (1 gene/5.4 kb) in the *Brassica campestris* S locus. Heteromorphic systems are also spread mechanisms but their functioning is quite different because only two "haplotypes" control mating. But such SI systems can also shelter deleterious or lethal alleles as has been found in the S "supergene" of Primula species (*e.g.*, Kurian and Richards 1997).

Despite these limitations, the more general result of this work is that interactions between loci under selection can significantly modify expectations of the levels of inbreeding depression and mutation load. In particular, it is striking that a few loci subject to strong balancing selection can significantly increase the load due to deleterious mutations.

S.G. thanks X. Vekemans, M. Uyenoyama, and D. Charlesworth for helpful discussions while this work was being done. We thank two anonymous reviewers for their constructive criticisms. This work was supported by the Bureau des Ressources Génétiques; by the French Ministère de L'Aménagement du Territoire et de l'Environnement through the national program Diversitas, Fragmented population network (contract 98/153 to I.O.); as well as by the EC Fragland project (headed by I. Hanski). This is publication ISEM 2001-072 of the Institut des Sciences de l'Evolution de Montpellier.

LITERATURE CITED

Aspinwall, N., and T. Christian, 1992 Pollination biology, seed production, and population structure in Queen-of-the-prairie, *Filipendula rubra* (Rosaceae) at Botkin fen, Missouri. Am. J. Bot. **79:** 488–494.

Awadalla, P., and D. Charlesworth, 1999 Recombination and selection at *Brassica* self-incompatibility loci. Genetics **152:** 413–425.

Baker, H. G., 1955 Self-compatibility and establishment after "long-distance" dispersal. Evolution **9:** 347–348.

Baker, H. G., 1967 Support for Baker's law as a rule. Evolution **21:** 853–856.

Barrett, S. C. H., M. T. Morgan and B. C. Husband, 1989 The dissolution of a complex genetic polymorphism: the evolution of self-fertilization in tristylous *Eichhornia paniculata* (Pontederiaceae). Evolution **43:** 1398–1416.

Bataillon, T., and M. Kirkpatrick, 2000 Inbreeding depression due to mildly deleterious mutations in finite populations: Size does matter. Genet. Res. **75:** 75–81.

Boyes, M., E. Nasrallah, J. Vrebalov and J. B. Nasrallah, 1997 Self-incompatibility (S) haplotypes of *Brassica* contain divergent and rearranged sequences of ancient origin. Plant Cell **9:** 237–247.

Byers, D. L., and T. R. Meagher, 1992 Mate availability in small populations of plant species with homomorphic sporophytic self-incompatibility. Heredity **68:** 353–359.

Casselman, A. L., J. Vrebalov, J. A. Conner, A. Singhal, J. Giovannoni *et al.*, 2000 Determining the physical limits of the *Brassica* S locus by recombination analysis. Plant Cell **12:** 23–33.

Charlesworth, B., M. Nordborg and D. Charlesworth, 1997 The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet. Res. **70:** 155–174.

Charlesworth, D., 1988 Evolution of homomorphic sporophytic self-incompatibility. Heredity **60:** 445–453.

Charlesworth, D., and B. Charlesworth, 1979 The evolution and breakdown of S-allele systems. Heredity **43:** 41–55.

Charlesworth, D., and B. Charlesworth, 1987 Inbreeding depression and its evolutionary consequences. Annu. Rev. Ecol. Syst. **18:** 237–268.

Charlesworth, D., M. T. Morgan and B. Charlesworth, 1990 Inbreeding depression, genetic load, and the evolution of out-

crossing rates in a multilocus system with no linkage. Evolution **44:** 1469–1489.
Charlesworth, D., M. T. Morgan and B. Charlesworth, 1992 The effect of linkage and population size on inbreeding depression due to mutational load. Genet. Res. **59:** 49–61.
Colas, B., I. Olivieri and M. Riba, 1997 *Centaurea corymbosa*, a cliff-dwelling species tottering on the brink of extinction: a demographic and genetic study. Proc. Natl. Acad. Sci. USA **94:** 3471–3476.
Crow, J. F., and M. Kimura, 1970 *An Introduction to Population Genetics Theory*, Alpha edition. Burgess, Minneapolis.
De Nettancourt, D., 1997 Incompatibility in angiosperms. Sex. Plant Reprod. **10:** 185–199.
Dowd, P. E., A. G. McCubbin, X. Wang, J. A. Verica, T. Tsukamoto *et al.*, 2000 Use of *Petunia inflata* as a model for the study of Solanaceous type self-incompatibility. Ann. Bot. **85:** 87–93.
Entani, T., M. Iwano, H. Shiba, S. Takayama, K. Fukui *et al.* 1999 Centromeric localization of an S-RNase gene in *Petunia hybrida* Vilm. Theor. Appl. Genet. **99:** 391–397.
Fisher, R. A., 1941 Average excess and average effect of a gene substitution. Ann. Eugen. **11:** 52–63.
Gale, J. S., 1990 *Theoretical Population Genetics*. Unwin Hyman, London.
Goodwillie, C., 1999 Multiple origins of self-compatibility in *Linanthus* section *Leptosiphon* (Polemoniaceae): phylogenetic evidence from internal-transcribed-spacer sequence data. Evolution **53:** 1387–1395.
Johnston, M. O., and D. J. Schoen, 1995 Mutation rates and dominance levels of genes affecting total fitness in two Angiosperm species. Science **267:** 226–229.
Kao, T., and A. G. McCubbin, 1996 How flowering plants discriminate between self and non-self pollen to prevent inbreeding. Proc. Natl. Acad. Sci. USA **93:** 12059–12065.
Karron, J. D., 1987 A comparison of levels of genetic polymorphism and self-incompatibility in geographically restricted and widespread plant congeners. Evol. Ecol. **1:** 47–58.
Kimura, M., and J. F. Crow, 1964 The number of alleles that can be maintained in a finite population. Genetics **49:** 725–738.
Kimura, M., T. Maruyama and J. F. Crow, 1963 The mutation load in small populations. Genetics **48:** 1303–1312.
Kurian, V., and A. J. Richards, 1997 A new recombinant in the heteromorphy 'S' supergene in *Primula*. Heredity **78:** 383–390.
Lane, M. D., and M. J. Lawrence, 1995 The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. X. An association between incompatibility genotype and seed dormancy. Heredity **75:** 92–97.
McCubbin, A. G., and T. H. Kao, 1999 The emerging complexity of self-incompatibility (S-) loci. Sex. Plant Reprod. **12:** 1–5.
Ohta, T., and M. Kimura, 1969 Linkage disequilibrium due to random genetic drift. Genet. Res. **13:** 47–55.
Ohta, T., and M. Kimura, 1970 Development of associative overdominance through linkage disequilibrium in finite populations. Genet. Res. **16:** 165–177.
Pillen, K., O. Pineda, C. B. Lewis and S. D. Tanksley, 1996 Status of genome mapping tools in the taxon solanaceae, pp. 281–308 in *Genome Mapping in Plants*, edited by A. H. Paterson. R. G. Landes, Austin, TX.
Reinartz, J. A., and D. H. Les, 1994 Bottleneck-induced dissolution of self-incompatibility and breeding system consequences in *Aster furcatus* (Asteraceae). Am. J. Bot. **81:** 446–455.
Schierup, M. H., X. Vekemans and F. B. Christiansen, 1997 Evolutionary dynamics of sporophytic self-incompatibility alleles in plants. Genetics **147:** 835–846.
Schierup, M. H., D. Charlesworth and X. Vekemans, 2000 The effect of hitch-hiking on genes linked to a balanced polymorphism in a subdivided population. Genet. Res. **76:** 63–73.
Simmons, M., and J. F. Crow, 1977 Mutations affecting fitness in *Drosophila* populations. Annu. Rev. Genet. **11:** 49–78.
Slatkin, M., 1991 Inbreeding coefficients and coalescence times. Genet. Res. **58:** 167–175.
Strobeck, C., 1980 Heterozygosity of a neutral locus linked to a self-incompatibility locus or a balanced lethal. Evolution **34:** 779–788.
Strobeck, C., 1983 Expected linkage disequilibrium for a neutral locus linked to a chromosomal arrangement. Genetics **103:** 545–555.
Susuki, G., N. Kai, T. Hirose, K. Fukui, T. Nishio *et al.*, 1999 Genomic organization of S locus: identification and characterization of genes in SLG/SRK region of S haplotype of *Brassica campestris*. Genetics **153:** 391–400.
Takahata, N., 1990 A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. Proc. Natl. Acad. Sci. USA **87:** 2419–2423.
Takahata, N., and Y. Satta, 1998 Footprints of intragenic recombination at HLA loci. Immunogenetics **47:** 430–441.
Uyenoyama, M. K., 1988a On the evolution of genetic incompatibility systems. II. Initial increase of strong gametophytic self-incompatibility under partial selfing and half-sib mating. Am. Nat. **131:** 700–722.
Uyenoyama, M. K., 1988b On the evolution of genetic incompatibility systems. I. Incompatibility as a mechanism for the regulation of outcrossing distance, pp. 212–232 in *The Evolution of Sex: An Examination of Current Ideas*, edited by M. Levin. Sinauer Associates, Sunderland, MA.
Uyenoyama, M. K., 1988c On the evolution of genetic incompatibility systems. III. Introduction of weak gametophytic self-incompatibility under partial inbreeding. Theor. Popul. Biol. **34:** 47–91.
Uyenoyama, M. K., 1989 On the evolution of genetic incompatibility systems. V. Origin of sporophytic self-incompatibility in response to overdominance in viability. Theor. Popul. Biol. **36:** 339–365.
Uyenoyama, M. K., 1991 On the evolution of genetic incompatibility systems. VI. A three-locus modifier model for the origin of gametophytic self-incompatibility. Genetics **128:** 453–469.
Uyenoyama, M. K., 1997 Genealogical structure among alleles regulating self-incompatibility in natural populations of flowering plants. Genetics **147:** 1389–1400.
Vekemans, X., and M. Slatkin, 1994 Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics **137:** 1157–1165.
Vekemans, X., M. H. Schierup and F. B. Christiansen, 1998 Mate availability and fecundity selection in multi-allelic self-incompatibility systems in plants. Evolution **52:** 19–29.
Weller, S. G., M. J. Donoghue and D. Charlesworth, 1995 The evolution of self-incompatibility in flowering plants: a phylogenetic approach, pp. 355–382 in *Experimental and Molecular Approaches to Plant Biosystematics*, edited by P. C. Hoch and A. G. Stephenson. Missouri Botanical Garden, St Louis.
Whishler, S. L., and A. A. Snow, 1992 Potential for the loss of self-incompatibility in pollen-limited populations of mayapple (*Podophyllum peltatum*). Am. J. Bot. **79:** 1273–1278.
Willis, J. H., 1999 Inbreeding load, average dominance and the mutation rate for mildly deleterious alleles in *Mimulus guttatus*. Genetics **153:** 1885–1898.
Wright, S., 1939 The distribution of self-sterility alleles in populations. Genetics **24:** 538–552.
Yokoyama, S., and L. E. Hetherington, 1982 The expected number of self-incompatibility alleles in finite plant populations. Heredity **48:** 299–303.
Yokoyama, S., and M. Nei, 1979 Population dynamics of sex-determining alleles in honey bees and self-incompatibility alleles in plants. Genetics **91:** 609–626.

Communicating editor: M. K. Uyenoyama

APPENDIX A

To derive F_{IS} for a neutral locus linked to the S locus, we use a coalescent model that was developed by Takahata and Satta (1998) to analyze nucleotide diversity in a neutral region partially linked to an HLA locus under strong balancing selection. They compute the mean coalescence time of two neutral genes that are sampled from the same (T_w) or different (T_b) overdominant allelic lineages (see the appendix of their article),

$$T_w = \frac{2N(\rho + r^*n)}{(n + 2N\rho)(\rho + r^*)}$$

and

$$T_b = T_w + \frac{n-1}{2(\rho + r^*)},$$

where n is the number of common overdominant alleles; $r^* = r(1 - 1/n)$, where r is the recombination rate between the two loci; and ρ is the rate of allelic turnover per lineage (see Takahata 1990). For a locus under balancing selection, common alleles are regularly replaced by initially rare alleles. r corresponds to the average turnover of alleles divided by the number of common alleles. They also compute the mean coalescence time of two genes randomly sampled from the population (regardless of allelic lineage):

$$T = \frac{1}{n} T_w + \left(1 - \frac{1}{n}\right) T_b.$$

These computations can be applied in our case because gene and allelic genealogies at a GSI locus behave in a similar way (Vekemans and Slatkin 1994). Moreover, F-statistics can be defined using coalescence times (Slatkin 1991) and so

$$F_{IS} = \frac{T_1 - T_0}{T_1},$$

where T_1 is the mean coalescence time of two genes randomly sampled from a population and T_0 is the mean coalescence time of two genes sampled in the same individual. Because all individuals are heterozygous at the S locus, we have $T_0 = T_b$ (and $T_1 = T$). So, using expressions for T_b and T, we obtain

$$F_{IS} = -\frac{(n-1)(n + 2N\rho)}{n(n-1)^2 + 2N\rho(1 + n^2) + 4Nnr(n-1)}.$$

ρ is directly related to the scaling factor, f_S, of the genealogy of overdominant or S alleles, which is a more useful parameter, always >1 (Takahata 1990; Vekemans and Slatkin 1994): $\rho = n/4Nf_S$, so we can express F_{IS} as

$$F_{IS} = -\frac{(1 + 2f_S)(n-1)}{1 + n^2 + 2f_S(n-1)(4Nr + n - 1)}$$

$$\approx -\frac{2f_S(n-1)}{n^2 + 2f_S(n-1)(4Nr + n - 1)}$$

for $f_S \gg 1$ and $n^2 \gg 1$. F_{IS} is always negative and is an increasing function of r.

The number of common alleles and the scaling factor can be determined following Vekemans and Slatkin (1994). The number of common alleles is approximately equal to the effective number of alleles: $n \approx n_e = 1/J$, where $J = \Sigma x_i^2$ is the expected homozygosity at the S locus (x_i is the frequency of the ith allele). The scaling factor is given by

$$f_S = \frac{\sqrt{2}}{16N^2 \mu_S J(J - \mu_S/\alpha)^2},$$

where μ_S is the mutation rate at the S locus (infinite allele model) and $\alpha = 1/(1 - J)(1 - 2J)$ (Yokoyama and Nei 1979). J can only be obtained numerically as the solution of

$$\mu_S \sqrt{8\pi N}\, e^{2NJ/(1-J)(1-2J)} = (1 - J)^{-1/2}(1 - 2J)^{-N[1/(1-J) + 2\mu_S]}$$

(Yokoyama and Hetherington 1982).

For a large range of parameters' values (especially for small μ_S and N), $2f_S(n-1)^2 \gg n^2$ and $2f_S(n-1)(4Nr + n - 1) \geq 2f_S(n-1)^2$ so $2f_S(n-1)(4Nr + n - 1) \geq n^2$. F_{IS} thus can be approximated by $F_{IS} \approx -1/(n - 1 + 4Nr)$.

APPENDIX B

We consider that total inbreeding depression (δ) is due to p loci all over the genome, each causing the same elementary inbreeding depression (d_1). We suppose that these loci act multiplicatively. In a panmictic population,

$$\delta = 1 - (1 - d_1)^p.$$

Expression of GSI significantly modifies inbreeding depression only if some loci are linked to the S locus. Assume that there are k such loci and $p - k$ unlinked loci, spread over the genome. Each linked locus causes the same elementary inbreeding depression (d_S), and each unlinked locus causes the depression d_2. The total inbreeding depression is

$$\delta_{GSI}(k) = 1 - (1 - d_2)^{p-k}(1 - d_S)^k.$$

We then compute δ and $\delta_{GSI}(k)$ as a function of the population size, using numerical values of Table 2. For δ, we use values from the panmixia column. For $\delta_{GSI}(k)$, we use values from the $r = 0.5$ column for d_2, and values from $r = 10^{-4}$ for d_S. As $\mu_1 = 10^{-3}$, we assume $p = 500$ so that the genomic mutation rate toward deleterious mutation, U, is equal to 0.5.